EDUSCHOOL INTERNATIONAL JOURNAL OF DATA SCIENCE AND MACHINE LEARNING (EIJDSML)

(Open Access, Double-Blind Peer Reviewed Journal)

ISSN Online: 3139-7344




# Enhancing Cardiac Auscultation with Machine Learning and Deep Learning: A Survey on Heart Sound Analysis

Akhil K K[1*], R Devi[2]

[1] Research Scholar, Department of Computer Science, School of Computer Science, Vels Institute of Science Technology & Advanced Studies, Pallavaram, India
[2] Professor & Head, Department of Applied Computing & Emerging Technologies, Vels Institute of Science Technology & Advanced Studies, Pallavaram, India

**Article information**

Received: 10th September 2026
Received in revised form: 1st October 2026
Accepted: 10th October 2026
Available online: 18th October 2026

Volume: 1
Issue: 3
DOI: https://doi.org/10.63090/EIJDSML/3139.7344.0013

**Abstract**

Cardiovascular disease kills more people than anything else. Early detection changes outcomes, and heart sounds are one of the cheapest places to look for the first signs. Phonocardiogram (PCG) recordings carry information about valve function that a trained ear can pick up, which is why auscultation has survived two centuries of newer imaging and why it is still the first test a patient meets. But listening is slow. It needs experience, and two clinicians can disagree about the same murmur. Machine learning (ML) and deep learning (DL) offer a way to automate that step and make it reproducible. This review asks what the literature has actually achieved, and what it has not. We searched IEEE Xplore, SpringerLink, PubMed, ScienceDirect and Google Scholar for work published between 2015 and 2024. Section 2 gives the search strings, the screening steps and the inclusion rules. Thirty-seven primary studies survived. Each was read in full, and its dataset, signal processing chain, model and reported performance were recorded in a comparison table that covers thirteen of them in detail. Reported accuracy in this field is high and rising. It is also, taken at face value, misleading. The studies reviewed here use different corpora, different class definitions, different segmentation strategies and different validation protocols, so 99.80% in one paper and 73.6% in another are not measuring the same thing. Standardised datasets are missing. So are agreed metrics. Paediatric recordings are thin on the ground, recording conditions vary, and almost nothing here has been tested in the clinic where it would have to work. The review pulls recent advances together, then sets the reported numbers against the methodological limits that decide how much those numbers are worth, which surveys in this area tend to skip.

**Keywords**:- Cardiovascular diseases, Machine Learning, Deep learning, Heart sound classification, Phonocardiogram

## I. INTRODUCTION

Heart disease remains the largest single cause of death worldwide. Much of it announces itself in sound. When a valve leaks or narrows, blood moving through it turns turbulent, and that turbulence reaches the chest wall as a murmur [1] [6]. Clinicians have used murmurs as a first indicator of valve trouble for as long as the stethoscope has existed [3] [8] [38]. Grading them is another matter. Quality, timing and intensity are all judged by ear, and that judgement is itself treated as a diagnostic feature in cardiovascular assessment [2] [7].

Auscultation survives because it is cheap, fast and needs no more equipment than a stethoscope. It is also hard. Heart sounds sit low in frequency, they are quiet, and they compete with breathing, gut noise and whatever else the patient's body is doing at the time [4]. Accuracy tracks the listener's experience closely. The examiner's skill with the instrument, the room, and the character of the sound itself all shape what ends up written in the notes [9] [10].

Those limits are what pushed AI into the problem. A murmur can point to valvular disease, to a congenital defect, or to heart failure, and telling those apart by ear is a skill that takes years to build and stays uneven even then. ML and DL models can be trained to make the same call from a raw recording. They make it in seconds [5].

None of this is free. These methods depend on how the audio was captured, on what the signal processing chain does to it before the model ever sees it, and on having enough well-labelled recordings to learn from. Picking a model that separates normal from abnormal reliably is a further problem, and not a solved one.

This review works through recent progress in heart sound analysis and asks how far ML and DL have moved past the limits of manual auscultation. Signal processing, datasets and classification methods each get a section. The remaining challenges are named. So are the directions that look most likely to help, and the limits on what the reported numbers can support.

Section 2 covers the search. Section 3 is the body of the review, and Section 4 collects the challenges and where they point. Section 5 closes.

## II. SEARCH STRATEGY

### A. Research questions

Four questions shaped the search and, later, the way the findings are organised.

- What are the systems used for capturing heart sounds?
- What are the signal processing techniques used for analyzing heart sounds?
- Which datasets are commonly used for training machine learning models in heart disease and murmur prediction?
- What classification techniques are applied to heart sound data for detecting murmurs and predicting heart diseases?

The four questions do double duty. They set the boundary of what was searched for, and in Section 3.5 they come back as the frame for what the corpus actually says.

### B. Data Collection Process

The search covered IEEE Xplore, PubMed, ScienceDirect, SpringerLink and Google Scholar, filtered to 2015-2024. Four query strings were used: ("heart sound" AND classification AND "machine learning"); ("heart murmur" AND detection AND ("artificial intelligence" OR "deep learning")); ("heart disease" AND prediction AND dataset); and ("machine learning" AND "heart disease"). Titles and abstracts came first. Full texts followed for anything that survived, and the reference lists of the retained papers were then checked by hand for work the strings had missed.

### C. Screening and Selection

Thirty-seven primary studies came through screening. Each was checked against the inclusion rules, with weight given to whether the method was described clearly enough to be judged, which dataset was used, and what was actually measured. A few further references were kept for the benchmark datasets and for background, not as primary studies. Three of the retained items are preprints ([16], [18] and [34]). They document systems or corpora the field already uses, which is why they are here, but their results have not been through peer review. One more thing about scope. Per-database yield counts are not reported, and this is a structured narrative review rather than a PRISMA-registered systematic review.

### D. Inclusion & Exclusion Criteria

Inclusion criteria:

- Articles published from 2015 to 2024 on heart disease and murmur prediction using ML and DL.
- Studies with empirical records or theoretical models focused on heart sound analysis and machine learning applications.
- Papers addressing key challenges like signal processing, classification accuracy, and real-time prediction.
- Peer-reviewed, open-access research papers, plus the preprints named in Section 2.3.
- Exclusion Criteria:

- Non-peer-reviewed studies, other than the preprints named in Section 2.3, and papers unrelated to heart disease or murmur prediction.
- Papers lacking sufficient methodological detail or those focusing on unrelated topics.

## III. ADVANCEMENTS IN HEART SOUND ANALYSIS: ML, DL AND EMERGING TRENDS

Demand for cardiac screening is rising faster than the supply of people trained to do it. That gap is the case for automating the listening step. Murmur sounds carry enough information to flag disease, and ML and DL methods have become the standard way of pulling that information out of a recording [11]. The analysis has sharpened. Fig. 1 shows the workflow most of this work follows, from capture through preprocessing and feature extraction to a classifier, and the subsections below take it stage by stage.

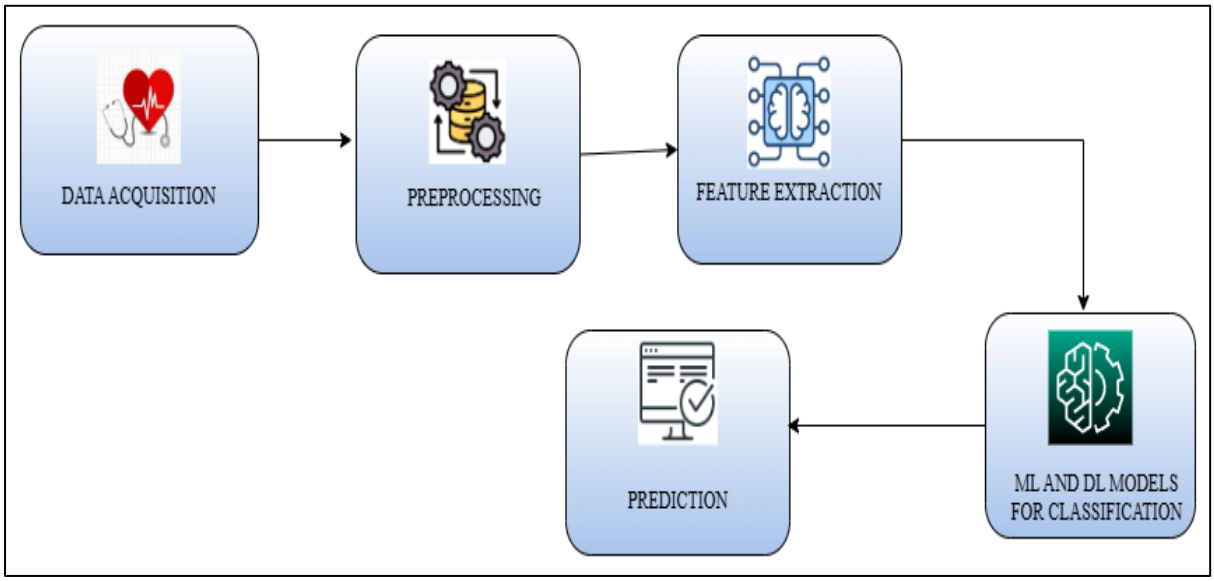



Fig 1: Heart Sound Analysis using ML and DL Models

### A. DATA ACQUISITION

1. Approaches for Capturing Heart Sounds

Capture hardware varies more than you would expect. The cheapest systems start from a stethoscope and a Bluetooth link. In [33] a stethoscope bell is fixed to the microphone of a dismantled Bluetooth headset, and [32] builds a Bluetooth sensing unit that passes recordings on for analysis. The system in [13] instead puts ECG and PCG amplification on a custom CMOS chip and shows both signals on a smartphone. The rig in [3] is heavier. Electrical detectors, analogue conditioning circuitry, a data acquisition card, a host computer and software for visualising the signal all sit in the chain.

The design in [15] is an electronic stethoscope assembled from parts that could be bought separately: a microphone, a preamplifier, a Butterworth filter, an AVR ATmega-32 microcontroller, a power amplifier and an LCD. Nothing exotic. A different approach appears in [4], where the sensor is worn as a finger ring and collects heart sounds alongside photoplethysmography (PPG).

Digital stethoscopes are the usual source of PCG recordings [16] [17]. Study [18] went after model reliability rather than raw accuracy, pairing Monte Carlo dropout for uncertainty estimation with temperature scaling for calibration. In [19], neuromorphic sensors that mimic the cochlea were mapped onto FPGAs so the processing could keep up with the signal in real time. Hardware is only half of it. The BUET multi-disease heart sound dataset [34] and the CirCor DigiScope dataset [35] both widened the range of recordings the field can train on, and public corpora of that kind have done as much for progress here as any sensor design.

2. Datasets Utilized

Fig. 2 counts how often each dataset appears across the reviewed studies. PhysioNet/CinC Challenge 2016 dominates with nine uses. CirCor DigiScope and the 2022 Challenge account for three each, MITHSDB for two, and BMD-HS, the paediatric corpus, Yaseen [36], Kaggle Heartbeat Sounds and PASCAL for one apiece. Two cautions apply to that tally. The categories overlap, because CirCor DigiScope is the corpus released for the 2022 Challenge and MITHSDB is one of the databases folded into the 2016 set, so the bars are not mutually exclusive. The Yaseen count also reads low against Table 1, where two of the tabulated studies use it.

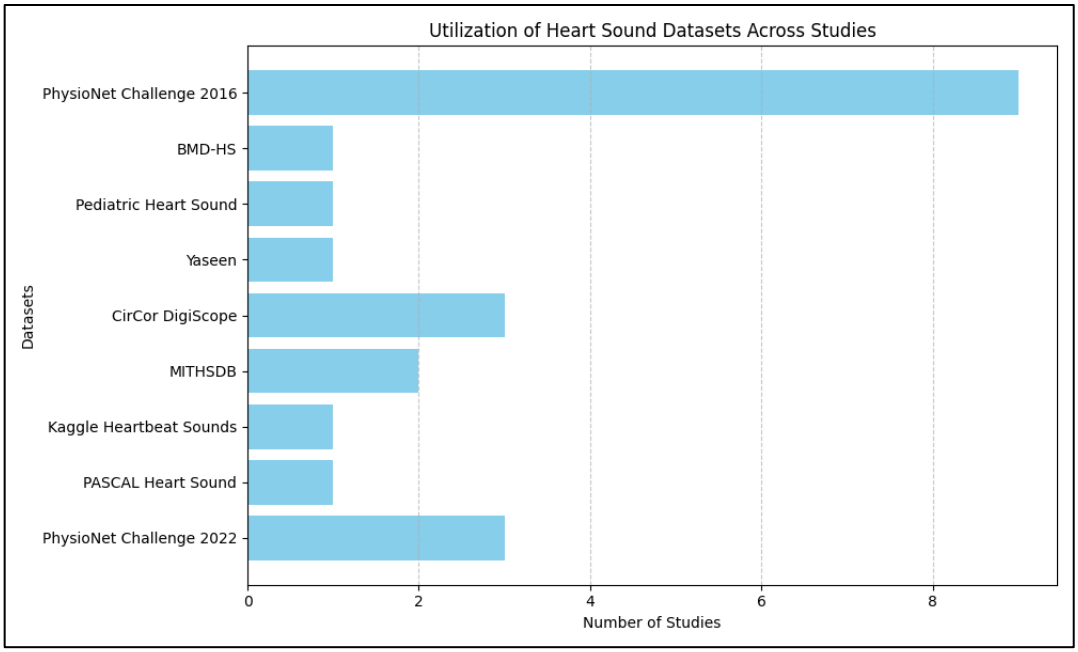


Fig 2: Dataset use across the reviewed studies

**B. Preprocessing**

Preprocessing decides how much signal survives to the model. Several chains have been built for it, and the differences between them matter more than their length suggests. The work in [1] [13] runs ECG and PCG through dual-channel analogue front ends, with amplifiers and a high-pass sigma-delta modulator doing the conditioning. Noise drops. The two streams stay synchronised, which is what makes a combined reading worth trusting.

A Bluetooth-linked acquisition unit appears in [32] [33]. Denoising there is handled by the Hilbert-Huang transform, while a double-threshold segmentation algorithm marks the physiological events, which leaves a cleaner trace with the cardiac events already located in it.

Shannon entropy has been used to find heart sound boundaries and photoplethysmography peaks [4] [8]. The intelligent stethoscope in [21] locks its ECG channel to the heart sound channel, and that synchronisation is what lets the two be read together rather than side by side. Small engineering choices carry weight here.

**C. Feature Extraction**

Feature extraction sets the ceiling. Whatever the features fail to encode, no downstream model recovers, which makes this stage worth more attention than the architecture choices that follow it. The Hilbert-Huang transform, mentioned above for denoising, doubles as a feature extractor, with double-threshold segmentation supporting the physiological analysis [32].

Continuous and discrete wavelet transforms carry the load in [16] and [7], where the problem is picking murmurs out of a noisy band. In [22] a denoising autoencoder learns deep features that feed a one-dimensional CNN; those features held up better under interference than Mel-frequency cepstral coefficients (MFCCs), and the improved network reached 99.01% accuracy. Learned features can beat engineered ones.

Decomposition methods do similar work elsewhere. Harmonic-percussive source separation, grey-level co-occurrence matrices and wavelet transforms all appear in [23] and [24]. Studies [25] and [26] convert PCG signals into two-dimensional images, chaogram and scalogram respectively, so that image models can be brought to bear. RNNs and hidden semi-Markov models split the signal into S1, S2, systole and diastole in [27]. The Temporal-Framing Adaptive Network in [29] does the same job without any prior assumption about how long each state lasts. Trained on only 50 recordings from the 2016 Challenge and tested on 2,099 more from twelve other databases, it reached an average F1 of 96.72%, ahead of both the logistic-regression HSMM and a bidirectional gated recurrent network. MFCCs go in directly as classifier inputs in [28], while [20] computes an improved MFCC variant and passes it to a deep residual network, reporting 94.43% accuracy. The scheme in [39] is the odd one out: short-term features computed over five-second windows are merged with long-term features taken from the whole recording, and across the twenty-two algorithms tested the merge lifted accuracy by 3 to 10%, with the best ensemble reaching 92.7%.

**D. Classification Using Ml Models**

CNNs do most of the work in this literature. The architecture in [21] separates heart sound signals by reading the spatial structure of their time-frequency representation, which is the same thing a vision model does to a photograph. AlexNet, transplanted whole from image recognition, classifies cardiac signals well enough to be clinically interesting [19]. That transplant is the field's most reused trick.

Hybrids are the current direction of travel. VGG16, ResNet50 and a Transformer encoder were each evaluated in [30] before the authors settled on a hybrid that feeds ResNet50 features into a Transformer encoder and an MLP head. Study [17] took a different route, mapping PCG signals into the time-frequency domain with the Stockwell transform, extracting deep features from those maps with AlexNet, and handing them to a support vector machine. It reports 93% accuracy, 91% sensitivity and an F1 of 91% on murmur detection.

CNN and LSTM layers are combined in [31], the CNN handling spatial structure and the LSTM the temporal. It is a natural fit. Biomedical signals are structured in both directions at once, and an end-to-end version of the same idea appears in [6], where segmentation and classification are learned together rather than staged one after the other.

Capsule networks and Bayesian neural networks address narrower problems. The CapsNet in [28] classifies murmurs by holding on to the spatial relationships and hierarchy that pooling would otherwise throw away, while a Dual Bayesian ResNet in [5] folds uncertainty estimation into the decision itself, so the model can report how confident it is and not only what it thinks. Study [2] fuses CNNs with channel attention and gated recurrent units, gaining on interpretability as well as performance. The Transformer in [40] sorts signals into four valve conditions plus normal. Per-class sensitivities above 98% are reported.

Classical classifiers have not disappeared. Zeinali and Niaki [14] select hand-crafted features and separate normal sounds from S3 and S4 with 87.5% to 95% accuracy on three classes and 98% on the binary task. Singh and Majumder [37] skip segmentation entirely: wavelet, Hilbert, homomorphic and spectral-density features from the first five seconds of each recording go to a k-nearest-neighbour classifier, which reaches 90% accuracy over the pooled datasets, with 93% sensitivity and 90% specificity. Neither needs a GPU.

Table 1 sets out the datasets, models and reported results for the studies discussed above. It needs reading with care. The figures come from different corpora, different class definitions, different segmentation strategies and different validation protocols, so 99.80% in one row and 73.6% in another are not comparable quantities; each says what one system achieved under its own protocol. Nothing in the table is a ranking. It is also worth saying plainly that every system here is a screening or decision-support aid, and none has been validated as a substitute for clinical assessment. The nearest thing to a field test in the corpus is [12]. There, a recurrent network trained on recordings from 2,124 participants of the Tromsø7 population study detected aortic stenosis with an AUC of 0.979 against echocardiography, yet picked up moderate or worse aortic and mitral regurgitation with AUCs of only 0.634 and 0.549.

Table 1. Datasets, models and reported results for the reviewed studies

| Authors | Datasets used | ML and DL models | Results |
|---|---|---|---|
| Wu, Tingwei, et al.[2] | • CirCor<br>• DigiScope Phonocardiogram dataset, | CNNs with channel attention and GRU networks | • **Accuracy = 73.6% (patient level)**<br>• **F1-scores = 76.8% (harsh murmurs), 67.8% (blowing murmurs)** |
| Walker, Benjamin, et al.[5] | George B. Moody PhysioNet Challenge 2022 | • Dual Bayesian ResNet (DBRes)<br>• XGBoost | • Weighted accuracy = 0.771 (hidden test set, 4th in the murmur task)<br>• Held-out accuracy = 76.2% (DBRes), 82.0% with XGBoost |
| Dominguez-Morales, Juan P., et al.[19] | Heart sound recordings from **nine different heart sound databases** sourced from multiple research groups. | • CNNs<br>• AlexNet model | • Accuracy: 97%<br>• Specificity=95.12%<br>• Sensitivity=93.20% |
| Manshadi, O. D., & Mihandoost, S. (2024)[17]. | **CinC/PhysioNet Challenge 2022** dataset | • Stockwell transform time-frequency mapsAlexNet deep featuresSVM classifier | • Accuracy=93%<br>• Sensitivity=91%<br>• F1 score=91% |

| | | | |
|---|---|---|---|
| Tsai, Yu-Ting, et al. [28] | • PhysioNet 2016 Heart Sound Database<br>• Clinical auscultation dataset | • CapsNet<br>• CNN<br>• MFCC | • Validation accuracy=90.29%<br>• Test accuracy=91.67% |
| Al-Tam, Riyadh M., et al. [30] | PhysioNet/CinC 2016 challenge datasetYaseen dataset | • Hybrid model(ResNet50+Transformer encoder=MLP) | Yaseen, multiclass: 99.80% accuracy and AUCYaseen, binary: 100% accuracy and AUCPhysioNet 2016, binary: 97.40% accuracy, 96.40% AUC |
| Yang, D., et al.[40] | Yaseen five-class dataset (AS, MR, MS, MVP, Normal) | • Transformer model | Per-class sensitivity (5-fold cross-validation):Aortic Stenosis: 98.72%Mitral Regurgitation: 98.50%Mitral Stenosis: 98.33%Mitral Valve Prolapse: 98.56%Normal: 99.61% |
| Arshad, Zainab, et al.[9] | • PhysioNet Heart Sound Database<br>• Cleveland Heart Disease Dataset<br>• Heart Sounds Dataset | • Naive Bayes<br>• (SMO)<br>• J48<br>• Rep Tree<br>• RF | • **Accuracy**=91.33% |
| Abbas, Sagheer, et al. [11] | • PASCAL challenge subsets + PhysioNet/CinC 2016 (832 signals, augmented to 2,882) | • MLP (compared against RF, KNN, DT, XGBoost, DNN, Conv1D) | • **Accuracy = 95.65%**<br>• **Precision = 96.60%**<br>• **Recall = 97.60%**<br>• **F1 = 96.60%**<br>• **(5 classes: normal, murmur, extrasystole, artifact, extra heart sound)** |
| Singh, Arora & Singh [23] | • PhysioNet/CinC 2016 (4,430 recordings, 1,072 patients) | • Harmonic-percussive spectral features + deep feed-forward ANN | • **All 82 harmonic features: Accuracy = 93.80%, Sensitivity = 82.50%, Specificity = 82.40%**<br>• **12 chroma features: Accuracy = 93.40%, AUC = 0.975** |
| Harimi, Alireza, et al. [25] | • PhysioNet/CinC (2,868 PCG signals: 2,249 normal, 619 abnormal) | • Chaogram transform + CNN transfer learning (AlexNet, VGG16, InceptionV3, ResNet50) | • **InceptionV3 (best): Accuracy = 90.36%, score = 88.06**<br>• **ResNet50 = 90.02%, VGG16 = 90.05%, AlexNet = 89.68%** |
| McDonald, Gales & Agarwal [27] | • CirCor DigiScope / PhysioNet-CinC 2022 (5,268 recordings, 1,452 patients) | • RNN + four parallel hidden semi-Markov models | • **Murmur detection, test set: weighted accuracy = 0.776, sensitivity = 92.7%**<br>• **Outcome task: AUROC = 0.691**<br>• **(First prize, both Challenge tasks)** |
| Guven & Uysal [39] | • PhysioNet/CinC 2016 (2,435 labelled recordings) | • Long short-term feature fusion + subspace KNN ensemble (22 algorithms compared) | • **Accuracy = 92.7% with merged long short-term features, against 90.3% for short-term features alone**<br>• **Gain across algorithms: 3-10%** |

### E. What the Corpus Says about the Four Questions

Capture (RQ1). Almost everything in the corpus starts from one of three things: a conventional stethoscope with a microphone bolted on, a purpose-built electronic stethoscope, or a wearable. Bluetooth is the default link. The interesting outliers are the finger-ring sensor in [4], which collects PPG alongside the heart sound, and the neuromorphic-plus-FPGA chain in [19], which pushes part of the processing into the sensor itself. Neither has been reproduced at scale.

Signal processing (RQ2). Wavelet transforms, the Hilbert-Huang transform, Shannon-entropy boundary detection and double-threshold segmentation cover most of the preprocessing seen here. Feature extraction has largely converged on two families. Cepstral coefficients, MFCCs above all, make up the first; two-dimensional time-frequency images that let pretrained vision models be reused make up the second. That move to images is what made transfer learning practical in this field.

Datasets (RQ3). The PhysioNet/CinC 2016 Challenge set is the field's default benchmark, with CirCor DigiScope, released for the 2022 Challenge, the main alternative, and the five-class Yaseen set [36] the usual choice for valve work. That concentration is a problem in its own right. When most of a field validates on two corpora, a reported gain partly measures fit to those corpora rather than to the condition.

Classification (RQ4). CNNs remain the backbone. What has changed since roughly 2022 is what gets attached to them: attention blocks, recurrent layers, Transformer encoders, capsule routing and Bayesian heads. The Bayesian and attention variants answer a clinical need rather than a benchmark one, because they return calibrated uncertainty [5] or an interpretable feature weighting [2] instead of a bare number.

## IV. CHALLENGES AND FUTURE DIRECTIONS

Heart sound analysis with ML and DL has gone from a niche interest to a crowded field in about a decade, and the sections above have traced how. Results range from modest to excellent. The obstacles standing between this work and a system anyone would actually deploy are more stubborn than the accuracy figures suggest. They are as follows:

- Lack of Standardised Evaluation Metrics and Protocols: reported accuracies are computed on different datasets, class definitions and train-test splits, and many are reported on imbalanced data without sensitivity, specificity or confidence intervals, which makes reliable benchmarking across studies difficult.
- Dataset Limitations: Most datasets focus on adult heart sounds, with limited representation of pediatric cases and diverse clinical conditions.
- Scarcity of Labeled Data: The shortage of annotated datasets restricts the training and evaluation of robust classification models.
- Underutilization of Multimodal Data: Combining PCG with other modalities, such as ECG or imaging, remains an underexplored avenue.
- Real-Time Detection Constraints: Hardware and resource limitations hinder the implementation of real-time detection systems, especially in low-resource settings.
- Preprocessing Challenges: Effective noise and artifact reduction methods are required to improve the quality of feature extraction from heart sound recordings.

The directions below follow from those six problems, roughly in the order the problems were listed.

- Develop and adopt standardized metrics to enable consistent benchmarking and fair comparisons across studies.
- Create datasets that include pediatric heart sounds and a variety of clinical conditions to improve model generalizability.
- Leverage transfer learning to maximize the use of limited labeled datasets and improve model performance in data-scarce scenarios.
- Integrate PCG signals with complementary data, such as ECG or medical imaging, for a more comprehensive analysis of cardiac health.
- Develop and optimize real-time detection systems using neuromorphic sensors and lightweight algorithms suitable for edge-based deployment.
- Refine preprocessing methods to effectively remove noise and artifacts, ensuring high-quality input data for classification models.

## V. CONCLUSION

This review has traced how ML and DL are being used to detect and classify heart murmurs, working from publications found through structured keyword searches across five databases. The methods were described and judged. Datasets and reported performance sit in Table 1. What follows from all of it, the open challenges and the directions they point to, was set out in Section 4.

Two things limit what has been said here. The review covers open-access work, which leaves subscription-based venues under-represented, and some of that literature is likely to matter. Performance figures were taken as published, not reproduced. Since they rest on different corpora and different protocols, they describe individual systems and not a league table. Read with those limits in mind, this review is meant to give anyone working on automated heart sound detection a single place to start and, further down the line, to help cardiac problems get caught earlier.

## REFERENCES

[1] M. E. H. Chowdhury, A. Khandakar, K. Alzoubi, S. Mansoor, A. M. Tahir, M. B. I. Reaz, and N. Al-Emadi, “Real-time smart-digital stethoscope system for heart diseases monitoring,” *Sensors*, vol. 19, no. 12, Art. no. 2781, 2019, doi: 10.3390/s19122781.

[2] T. Wu, Z. Huang, S. Li, Q. Zhao, and F. Pan, "Heart murmur quality detection using deep neural networks with attention mechanism," *Applied Sciences*, vol. 14, no. 15, Art. no. 6825, 2024, doi: 10.3390/app14156825.
[3] C.-J. She, X.-F. Cheng, and K. Wang, "Analysis of heart-sound characteristics during motion based on a graphic representation," *Sensors*, vol. 22, no. 1, Art. no. 181, 2021, doi: 10.3390/s22010181.
[4] S. Lee, Q. Wei, H. Park, Y. Na, D. Jeong, and H. Lim, "Development of a finger-ring-shaped hybrid smart stethoscope for automatic S1 and S2 heart sound identification," *Sensors*, vol. 21, no. 18, Art. no. 6294, 2021, doi: 10.3390/s21186294.
[5] B. Walker, F. Krones, I. Kiskin, G. Parsons, T. Lyons, and A. Mahdi, "Dual Bayesian ResNet: A deep learning approach to heart murmur detection," *Computing in Cardiology*, vol. 49, 2022, doi: 10.22489/CinC.2022.355.
[6] Y. Chen, Y. Sun, J. Lv, B. Jia, and X. Huang, "End-to-end heart sound segmentation using deep convolutional recurrent network," *Complex & Intelligent Systems*, vol. 7, no. 4, pp. 2103–2117, 2021, doi: 10.1007/s40747-021-00325-w.
[7] M. Nabih-Ali, E.-S. A. El-Dahshan, and A. S. Yahia, "Heart diseases diagnosis using intelligent algorithm based on PCG signal analysis," *Circuits and Systems*, vol. 8, no. 7, pp. 184–190, 2017, doi: 10.4236/cs.2017.87012.
[8] N. Shakhovska and I. Zagorodniy, "Classification of acoustic tones and cardiac murmurs based on digital signal analysis leveraging machine learning methods," *Computation*, vol. 12, no. 10, Art. no. 208, 2024, doi: 10.3390/computation12100208.
[9] Z. Arshad, S. Masood Bhatti, H. Tauseef, and A. Jaffar, "Heart sound analysis for abnormality detection," *Intelligent Automation & Soft Computing*, vol. 32, no. 2, pp. 1195–1205, 2022, doi: 10.32604/iasc.2022.022160.
[10] S. Kamepalli, B. S. Rao, and K. Venkata Krishna Kishore, "Multi-class classification and prediction of heart sounds using stacked LSTM to detect heart sound abnormalities," in *2022 3rd International Conference for Emerging Technology (INCET)*, IEEE, 2022.
[11] S. Abbas, S. Ojo, A. Al Hejaili, G. A. Sampedro, A. Almadhor, M. M. Zaidi, and N. Kryvinska, "Artificial intelligence framework for heart disease classification from audio signals," *Scientific Reports*, vol. 14, no. 1, Art. no. 3123, 2024, doi: 10.1038/s41598-024-53778-7.
[12] P. N. Waaler, H. Melbye, H. Schirmer, M. K. Johnsen, T. Donnem, J. Ravn, et al., "Algorithm for predicting valvular heart disease from heart sounds in an unselected cohort," *Frontiers in Cardiovascular Medicine*, vol. 10, Art. no. 1170804, 2023, doi: 10.3389/fcvm.2023.1170804.
[13] S.-Y. Lee, P.-W. Huang, J.-R. Chiou, C. Tsou, Y.-Y. Liao, and J.-Y. Chen, "Electrocardiogram and phonocardiogram monitoring system for cardiac auscultation," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 13, no. 6, pp. 1471–1482, 2019, doi: 10.1109/TBCAS.2019.2947694.
[14] Y. Zeinali and S. T. A. Niaki, "Heart sound classification using signal processing and machine learning algorithms," *Machine Learning with Applications*, vol. 7, Art. no. 100206, 2022, doi: 10.1016/j.mlwa.2021.100206.
[15] J. J. Seah, J. Zhao, D. Y. Wang, and H. P. Lee, "Review on the advancements of stethoscope types in chest auscultation," *Diagnostics*, vol. 13, no. 9, Art. no. 1545, 2023, doi: 10.3390/diagnostics13091545.
[16] M. Jobayer et al., "FunnelNet: An end-to-end deep learning framework to monitor digital heart murmur in real-time," *arXiv*, 2024. [Online]. Available: http://arxiv.org/abs/2405.09570.
[17] O. D. Manshadi and S. Mihandoost, "Murmur identification and outcome prediction in phonocardiograms using deep features based on Stockwell transform," *Scientific Reports*, vol. 14, no. 1, Art. no. 7592, 2024, doi: 10.1038/s41598-024-58274-6.
[18] J. Nie, R. Liu, B. Mahasseni, E. Azemi, and V. Mitra, "Model-driven heart rate estimation and heart murmur detection based on phonocardiogram," *arXiv*, 2024. [Online]. Available: http://arxiv.org/abs/2407.18424.
[19] J. P. Dominguez-Morales, A. F. Jimenez-Fernandez, M. J. Dominguez-Morales, and G. Jimenez-Moreno, "Deep neural networks for the recognition and classification of heart murmurs using neuromorphic auditory sensors," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 12, no. 1, pp. 24–34, 2018, doi: 10.1109/TBCAS.2017.2751545.
[20] F. Li, Z. Zhang, L. Wang, and W. Liu, "Heart sound classification based on improved mel-frequency spectral coefficients and deep residual learning," *Frontiers in Physiology*, vol. 13, Art. no. 1084420, 2022, doi: 10.3389/fphys.2022.1084420.
[21] S.-Y. Lee, P.-H. Su, Y.-T. Hsieh, S.-H. Huang, I.-P. Lee, and J.-Y. Chen, "Intelligent stethoscope system and diagnosis platform with synchronized heart sound and electrocardiogram signals," *IEEE Access*, pp. 1–1, 2023, doi: 10.1109/ACCESS.2023.3275015.
[22] F. Li, M. Liu, Y. Zhao, L. Kong, L. Dong, X. Liu, and M. Hui, "Feature extraction and classification of heart sound using 1D convolutional neural networks," *EURASIP Journal on Advances in Signal Processing*, vol. 2019, no. 1, 2019, doi: 10.1186/s13634-019-0651-3.
[23] Singh, V. Arora, and M. Singh, "Heart sound classification using harmonic and percussive spectral features from phonocardiograms with a deep ANN approach," *Applied Sciences*, vol. 14, no. 22, Art. no. 10201, 2024, doi: 10.3390/app142210201.
[24] V. Arora, E. Y.-K. Ng, R. S. Leekha, K. Verma, T. Gupta, and K. Srinivasan, "Health of things model for classifying human heart sound signals using co-occurrence matrix and spectrogram," *Journal of Mechanics in Medicine and Biology*, vol. 20, no. 6, Art. no. 2050040, 2020, doi: 10.1142/S0219519420500402.
[25] Harimi, Y. Majd, A. A. Gharahbagh, V. Hajihashemi, Z. Esmaileyan, J. J. M. Machado, and J. M. R. S. Tavares, "Classification of heart sounds using chaogram transform and deep convolutional neural network transfer learning," *Sensors*, vol. 22, no. 24, Art. no. 9569, 2022, doi: 10.3390/s22249569.

[26] M. Xiang, J. Zang, J. Wang, H. Wang, C. Zhou, R. Bi, et al., "Research of heart sound classification using two-dimensional features," *Biomedical Signal Processing and Control*, vol. 79, Art. no. 104190, 2023, doi: 10.1016/j.bspc.2022.104190.

[27] McDonald, M. J. F. Gales, and A. Agarwal, "A recurrent neural network and parallel hidden Markov model algorithm to segment and detect heart murmurs in phonocardiograms," *PLOS Digital Health*, vol. 3, no. 11, Art. no. e0000436, 2024, doi: 10.1371/journal.pdig.0000436.

[28] Y.-T. Tsai, Y.-H. Liu, Z.-W. Zheng, C.-C. Chen, and M.-C. Lin, "Heart murmur classification using a capsule neural network," *Bioengineering*, vol. 10, no. 11, Art. no. 1237, 2023, doi: 10.3390/bioengineering10111237.

[29] X. Wang, C. Liu, Y. Li, X. Cheng, J. Li, and G. D. Clifford, "Temporal-framing adaptive network for heart sound segmentation without prior knowledge of state duration," *IEEE Transactions on Biomedical Engineering*, vol. 68, no. 2, pp. 650–663, 2021, doi: 10.1109/TBME.2020.3010241.

[30] R. M. Al-Tam, A. M. Al-Hejri, E. Naji, F. A. Hashim, S. S. Alshamrani, A. Alshehri, and S. M. Narangale, "A hybrid framework of transformer encoder and residential conventional for cardiovascular disease recognition using heart sounds," *IEEE Access*, vol. 12, pp. 123099–123113, 2024, doi: 10.1109/ACCESS.2024.3451660.

[31] N. B. Aji, K. Kurnianingsih, N. Masuyama, and Y. Nojima, "CNN-LSTM for heartbeat sound classification," *JOIV: International Journal on Informatics Visualization*, vol. 8, no. 2, p. 735, 2024, doi: 10.62527/joiv.8.2.2115.

[32] Ren, H., Jin, H., Chen, C., Ghayvat, H., and Chen, W., "A novel cardiac auscultation monitoring system based on wireless sensing for healthcare," *IEEE*.